



# Epistemological and Methodological Eclecticism in the Construction of Knowledge Organization Systems (KOSs) The Case of Analytico-synthetic KOSs.

Thomas Dousa, Fidelia Ibekwe-Sanjuan

## ► To cite this version:

Thomas Dousa, Fidelia Ibekwe-Sanjuan. Epistemological and Methodological Eclecticism in the Construction of Knowledge Organization Systems (KOSs) The Case of Analytico-synthetic KOSs.. Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects, May 2014, Krakow, Poland. pp.152-159. hal-01066220

**HAL Id: hal-01066220**

**<https://hal.science/hal-01066220>**

Submitted on 19 Sep 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thomas M. Dousa (University of Illinois, Urbana-Champaign)  
Fidelia Ibekwe-SanJuan (ELICO - University of Lyon 3)

## Epistemological and Methodological Eclecticism in the Construction of Knowledge Organization Systems (KOSs) The Case of Analytico-synthetic KOSs

### Abstract

In recent years, Hjørland has developed a typology of basic epistemological approaches to KO that identifies four basic positions—empiricism, rationalism, historicism/hermeneutics, and pragmatism—with which to characterize the epistemological bases and methodological orientation of KOSs. Although scholars of KO have noted that the design of a single KOS may incorporate epistemological-methodological features from more than one of these approaches, studies of concrete examples of epistemologico-methodological eclecticism have been rare. In this paper, we consider the phenomenon of epistemologico-methodological eclecticism in one theoretically significant family of KOSs—namely analytico-synthetic, or faceted, KOSs—by examining two cases—Julius Otto Kaiser’s method of Systematic Indexing (SI) and Brian Vickery’s method of facet analysis (FA) for document classification. We show that both of these systems combined classical features of rationalism with elements of empiricism and pragmatism and argue that such eclecticism is the norm, rather than the exception, for such KOSs in general.

### Introduction

Epistemology is the area of philosophy that investigates how human beings (come to) know what they know and, as such, it seeks to give a theoretical account of what constitutes knowledge. In recent years, the epistemological foundations of classification and indexing have increasingly become a subject of interest within the field of knowledge organization (KO). Within the framework of an ambitious project to characterize and evaluate different approaches to research and practice in KO and, more generally, in the field of library and information science (LIS), Birger Hjørland has elaborated, in a number of publications (e.g., Hjørland, 1998, 2003, 101, 105–107; 2011; 2013a, 174, Table 1), a typology of epistemological positions that identifies four primary streams of philosophical views regarding knowledge: *empiricism*, which holds that knowledge is derived from the observation of phenomena and the drawing of inductions therefrom; *rationalism*, which posits that knowledge is formed through the rational intuition of general categories and the structuring of classes within those categories on the basis of rules of logical definition and division; *historicism*, or *hermeneutics*, which emphasizes the historical, developmental, and, generally, broader contextual factors involved in the shaping of knowledge; and *pragmatism*, or *critical theory*, which foregrounds the knower’s goals and values, as well as the intellectual and practical consequences of what he or she knows, as key moments in the constitution of knowledge.

According to Hjørland (2003, 105), “the methods of KO” are to be related to these four “fundamental theories of epistemology”. Thus, he has posited that there are “four basic kinds of methods” for KO—“empirical methods”, “rationalistic methods”, “historical methods”, and “pragmatic methods” (Hjørland, 2002, 259). On this view, the epistemological positions of empiricism, rationalism, historicism, and pragmatism underwrite the methods

of KO but are analytically distinct from the latter (Hjørland, 1998, 163). This opens up the possibility that the use of a given method within a given KO system can, in principle, be dissociated from the epistemological position providing its theoretical justification. An important corollary of this is that the methods used in designing a given KO system need not mirror, in all respects, the core epistemological commitments of its designer(s). Indeed, it has been observed that specific (KOSs) do not limit themselves exclusively to methods deriving from a single epistemological position but rather tend to draw upon methodological criteria culled from different epistemological viewpoints (Hjørland, 2003, 107). As Hjørland (2013a, 173) has recently put it, “[s]ome classifications are based on a mixture of criteria (e.g., combined logical, empirical, historicist, and pragmatic criteria”. In practice, then, the epistemologico-methodological bases for classification and indexing tend to be eclectic in nature.

Although theorists of KO have not failed to point out instances in which certain design features of KOSs involve a measure of epistemologico-methodological eclecticism (e.g., Tennis, 2008, 108), concrete and detailed analyses of such eclecticism in the creation of particular (kinds of) systems have been rather rare (Dousa, 2008; Ibekwe-SanJuan, in Fuller et al., 2013). The purpose of this paper is to broaden the evidentiary base for epistemologico-methodological eclecticism by examining one well-known family of KOSs—namely, those falling broadly within the analytico-synthetic tradition of facet analysis—through the prism of Hjørland’s typology. Now, in his writings, Hjørland (2003, 105 & 107, Figure 5; 2011, 74; 2013b) has characterized faceted classifications and indexing systems as being primarily rooted in rationalist assumptions. Here, we shall examine the epistemologico-methodological bases for the design protocols of two forms of analytico-synthetic KOSs—namely, the method of Systematic Indexing developed by Julius Otto Kaiser (1911) and the methodological guide-lines for constructing specialized faceted classifications enunciated by Brian C. Vickery (1958, 1960, 1966): in doing so, we intend to demonstrate that methods for designing analytico-synthetic KOSs tend to fuse elements of rationalist, empiricist, and pragmatist approaches in certain predictable ways.

### **Case 1: Julius Otto Kaiser’s Method of Systematic Indexing**

Let us begin with Kaiser and his indexing system. Julius Otto Kaiser (1868–1927) was a special librarian and indexing specialist who spent his career working in commercial and technical libraries in the United States of America and the United Kingdom (Sales, 2012, 49–54). Towards the end of the 19<sup>th</sup> century, he devised a highly analytical method for indexing commercial and technical literature that he dubbed Systematic Indexing (SI) (Kaiser, 1911, §§ 20–21). Envisioning that his system would be used primarily to construct customized indexes within libraries belonging to business organizations, Kaiser designed it for the purpose of indexing highly specific pieces of information *within* documents rather than documents as a whole: The aim was to “take literature to pieces and re-arrange the pieces systematically” (§ 21; cf. §§ 48, 298) within the framework of an alphabetically-arranged subject card index. Developed some thirty years before S. R. Ranganathan began to work out what would become an explicit theory of facet analysis, SI is today recognized

as a pioneering effort at applying an analytico-synthetic method to the organization of information (e.g., Sales, 2012, 16, 142, 153–155, 181).

A cardinal feature of SI was its use of a system of categories both as a means of dividing the terms comprising its index vocabulary into smaller, more tractable groups and as a mechanism for the structuring of complex index terms, which Kaiser (1911, §§ 301–302) dubbed “statements”. Each and every term forming part of an index vocabulary was to be assigned to one, and only one, of three categories: terms for *concretes*, which denoted things in general, with a special emphasis on commodities (§§ 52, 73, 299, 316); terms for *processes*, which referred to actions performed in relation to concretes or, more broadly, the various conditions pertaining to them (§§ 52, 73, 301); and terms for *countries*, which named the (geo)politically defined locales associated with concretes and their processes (§§ 300, 73, 302). Now Kaiser gave two explanations for his choice of these categories. On one hand, he held that concretes and processes represented the two fundamental—indeed, the only—kinds of objects of human knowledge in the world (§§ 52, 56), with countries constituting a special kind of concrete (§ 300). On the other, in accordance with the tenet that “literature”—that is to say textual documents—is a record of knowledge about the world (§§ 52–53, 297), he posited that the ontologico-epistemological distinction between concretes and processes was mirrored by a logico-linguistic distinction between things that are named (i.e., concretes and, by extension, countries) and that which is said about them (i.e., processes) in written texts (§§ 298, 301). For Kaiser, then, the categories of SI were not only derived from the structure of the world as it revealed itself to human observers (§§ 54–55) but also were “constant elements” in the texts that formed the object of indexing (§§ 298, 565–566, 574; cf. §§ 302, 663, “Analysis of Literature”): serving as the building blocks for the construction of complex subject statements, they formed, in his view, the “fixed points” upon which the entire edifice of his indexing system was built (§§ 574, 645).

Viewed from the perspective of Hjørland’s (2011, 74) typology, Kaiser’s use of a system of categories can be regarded as a classic example of a rationalist approach to indexing. Another side of SI had a more empiricist flavor. Kaiser (1911) stipulated that the terms incorporated into an index vocabulary were to be derived from the texts being indexed, either directly by extraction (§§ 114, 318, 320, 338, Point 4, 417) or, if necessary, indirectly by interpretative extrapolation from the text (§§ 322, 344, 346). As a rule, he favored term extraction—especially in the case of terms of concretes—because, as he saw it, authors choose the words they use in their writings to convey a certain meaning (§ 118) and he feared that substituting an author’s own words with different ones chosen by the indexer might result in a misrepresentation of the subject matter of the piece of information being indexed (§ 114; cf. § 67). To be sure, there were limits to Kaiser’s policy of derived indexing *avant la lettre*: he insisted that, once extracted from texts, terms were to be subjected to morphological normalization (§§ 319, 321–324, 329–330) and he allowed for the possibility that it might sometimes be necessary, for pragmatic reasons, to use a term other than the words found in the text (§§ 320, 328, 663, s. v. “Analysis of Literature”). Nevertheless, it is apparent that he assumed that, in most cases, the verbal formulations of the documents within which indexable pieces of information were identified would be the source of the

index terms indicating the subjects of those pieces of information: that is to say, he based, the choice of index terms primarily on what a later commentator would call “terminological warrant” (Beghtol, 1986, 113). Since terminological warrant is a form of literary warrant and literary warrant represents, in principle, an empirical approach to indexing (Svenonius, 2000, 135), this aspect of SI can be characterized as reflecting an empiricist imperative.

Alongside the rationalist and empiricist aspects of SI noted above, one may also discern elements of pragmatism. Here it is important to remind ourselves that Kaiser intended his indexing method to be used to construct indexes for libraries belonging to business organizations. In his eyes, such indexes had the function of supporting a business library’s mission of supplying businessmen with the information that they needed to conduct their business successfully (Kaiser, 1911, §§ 24, 36, 47–48). Sensitive to the fact that different businesses require information about different commodities and their conditions, Kaiser built into the protocols for SI measures to assure that the contents of any given index would be aligned with the interests of the business organization for which it was being constructed.

To take but one example, Kaiser assumed that any business would have its own particular sphere of interests. He thus posited that “[e]very firm or individual who has occasion to index literature does so ... to extract from it those parts which come within his purview” (Kaiser, 1911, § 309) or, to put it in more prescriptive terms, an indexer was to index only those pieces of information whose subjects fell within the sphere of interest of the business organization for which he or she worked. This meant that the subject scope of the index had to be carefully defined to reflect a business’s “particular sphere of activity” (§ 91). This, in turn, required that decisions be made regarding which subject terms to include within an index’s vocabulary, for, as Kaiser put it, “before [an indexer] starts his work, he must clearly define his purpose, he must determine on what concretes and processes information is wanted” (§ 311). Once the scope had been determined, it was incumbent upon the indexer to use it as a guide for the selection of information to be indexed. Thus, for instance, any given text was bound to mention a number of different concretes (e.g., §§ 448–449, 452) but an indexer was to select for indexing only those passages mentioning concretes that fell within the subject scope of his or her index (§§ 450, 453). Kaiser also held that any single passage of textual information might be indexed in different ways by indexers working for different kinds of business organizations: in his words, “each will index it [sci., the piece of information—TMD] from his own standpoint and for his own purpose” (§ 452). In strongly insisting that the subject analysis of pieces of information should be aligned to the goals, values, and activities of the organization for which an indexer was working, Kaiser exemplified an approach to indexing that can only be described, in Hjørland’s (2011, 74–75) terms, as pragmatist.

## **Case 2: Brian Vickery’s Facet Analysis (FA) for Document Classification**

Having seen that Kaiser’s method of SI combined elements of rationalism, empiricism, and pragmatism, we now turn to Vickery’s characterization of FA for document classification. Originally trained as a chemist, Brian Campbell Vickery (1919–2009) spent the early years of his career as a special librarian in science and technology libraries before taking up

research, administrative, and university teaching posts of various sorts (Broughton, 2011). In the 1950s, he was one of the leading lights of the British Classification Research Group (CRG), a body of librarians and documentalists who, influenced by S. R. Ranganathan's work on faceted classification, were at that time exploring the use of the techniques of faceted classification in constructing document classifications for special libraries (Broughton, 2011, 43; Vickery, 1966, 11). His experiences as a special librarian in a science library and as a member of the CRG informed a number of articles as well as monographs on classification in science and technology and on the principles and methodology of faceted classification that he wrote in the 1950s and 1960s (Vickery, 1958, 1960, 1966). These latter works, in the judgment of one recent commentator, "still stand as seminal texts in the field" of KO (Broughton, 2011, 44) and continue to serve as touchstones for present-day writers on FA (e.g., La Barre, 2010, 249–250). In short, Vickery's accounts of FA thus stand at the very heart of the facet-analytic tradition of KO.

"Facet analysis, by means of fundamental distinguishing characteristics or *categories*, is the basic operation in constructing a faceted classification" (Vickery, 1960, 13 [emphases his]). This lapidary statement neatly encapsulates the fundamental feature of FA as Vickery saw it. To his mind, the "essence" of FA lay in "the sorting of terms in a given field of knowledge into homogeneous, mutually exclusive facets, each derived from the parent universe by a single characteristic of division" (p. 12). On this view, *categories* were the "characteristics of division" used to partition the vocabulary of a field of study into *facets*, that is to say, semantically cohesive "groups of terms" in a manner that, according to Vickery, was "partly analogous to the traditional rules of logical division" (p. 12). Once terms had been grouped into facets, these could be further subdivided into classes and subclasses (pp. 38–39): as Vickery (1966, 31–32) put it, FA "first groups terms encountered in a given field into categories ... and then arranges the terms with [*sic*: read "within"] each category into a classificatory map". In addition to serving as a means of partitioning a universe of terms into facets, categories served as the building blocks for formulae governing the combination of single terms expressing simple subjects into compound ones expressing complex subjects, the sequence of simple terms within each compound term being determined by the category to which they belonged (1960, 13, 27–31). For the purpose of constructing classifications for specialized fields in science and technology, Vickery proposed no fewer than thirteen different general categories (p. 23) as well as principles for synthesizing compound subjects treated in documents written in scientific or technical fields (1960, 30–31; 1966, 54).

Vickery's use of categories and his invocation of logical division in relation to FA are, of course, hallmarks of a rationalist approach to knowledge organization (Hjørland, 2003, 105). Yet, perhaps unsurprisingly for somebody trained in the physical sciences, he also manifested an empiricist streak in his protocols for FA. For example, he held that, ideally, both the terms used to denote subjects in a classifications and the categories that served as the basis for dividing them into facets should be derived from "a detailed examination of the literature of the field to be classified" (Vickery, 1960, 20): to this end, he recommended that the designer of a classification peruse papers, reports, textbooks, glossaries, and other

documentary sources of information pertaining to the field in question (1966, 45). In the same vein, he cautioned that a prefabricated list of categories should not be “mechanically” (1960, 24) adopted in designing a classification:

[L]ists of categories have been developed by the study of various fiends of knowledge, and are undoubtedly helpful as a guide. But our grouping of terms in a newly approached field cannot be carried out *a priori*. New associations are continually being discovered and recorded in the literature, and we cannot too rigidly lay down in advance what *kinds* of term—i.e. what categories—will be associated. So in order to formulate categories, our first step must be a concrete examination of the literature of the given subject. Classification must be based—so far as its raw material is concerned—on literary warrant (1958, 23 [emphases his]).

Preexistent categories thus functioned, in Vickery’s eyes, as sensitizing concepts alerting the designer of a classification to possible classes for terms (1960, 24): however, only empirical engagement with the kinds of relations among terms actually found in texts on science or technology would give him or her the warrant to decide upon an appropriate set of categories for his or her classification.

Finally, one should not overlook the pragmatist elements embedded in Vickery’s account of FA for document classification. We have already noted that his methodological guidelines were intended primarily for the design of specialized classification schemes for particular scientific and technological fields. In his view, designing such a “special scheme” entailed taking into consideration “the field of interest of a particular group of users” and demarcating the subject scope of the classification in light of these interests (Vickery, 1966, 42–44). Inasmuch as the subject terms referred to entities and aspects thereof that were of interest to a particular community (pp. 43–44), the selection of terms and categories for facets would reflect the needs and interests of that community. Moreover, Vickery held that the choice of formulae for ordering simple terms within a compound subject term—which, in a printed or card-based classified catalog or index, had significant implications for retrieval—should take into account the expectations and interests of the prospective users: in his words,

[d]espite the utility of general combination formulae, it is unwise to rely entirely on them, however well tested they may appear to be. In a special classification, designed for specialist users, it is their interests that are paramount. ... Only sensitive attunement to the literary needs of his users can provide the specialist classifier with an answer to such problems (1960, 31; cf. pp. 34–35).

Interestingly, Vickery did not encourage the collection of “data on the grouping of subjects in the literature and on the approach of catalogue users”, since, in his view, “the systematic collection of such data for subjects containing a number of categories is a long and difficult task, impracticable though not impossible”, and analysis of the evidence collected was likely to lead to inconclusive results (p. 28). Rather, he noted, in making decisions about facet formulae, a classificationist might have to rely on “[a]n instinctive and empirical knowledge of helpful combination order” derived from his or her experience “in the course of work” (p. 28). This statement could, of course, be read as a more-or-less rationalist appeal to personal intuition, especially as Vickery went on to equate “instinctive and empirical knowledge” with “flair”. However, perhaps underlying it was the assumption that the designer of a classification scheme had immersed him or herself sufficiently in the thought-world of his or her intended user community to understand their goals, values, and

“literary needs” from the inside: otherwise, why the description of the designer’s knowledge as “empirical”? At any rate, here the line between rationalism and pragmatism seems to have been very thin indeed.

### Discussion and Concluding Comments

The foregoing analysis of Kaiser’s method of SI and Vickery’s method of FA in light of Hjørland’s typology is only partial and could readily be expanded. Nevertheless, it suffices to show that both of these methods for creating an analytico-synthetic KOS displayed epistemologico-methodological eclecticism and did so in similar ways. Kaiser and Vickery took a rationalist approach in deploying categories to partition the vocabulary of a KOS into smaller groupings and to construct complex subject statements; yet, they both took the empiricist principles of terminological or literary warrant as the basis for forming the vocabulary in the first place and both upheld—Kaiser more forcefully than Vickery—the pragmatist tenet that one should stock and structure the vocabulary of a KOS in a manner that corresponds to the interests, goals, and values of the user community for which it is being designed. In short, the use of rationalist structures was tempered by an empiricist attitude to vocabulary generation and a pragmatist concern with configuring structures and vocabulary to the interests of particular user groups.

It would perhaps be unwise to generalize too hastily this pattern to all faceted KOSs, for the particular conjunction of methodological rationalism, empiricism, and pragmatism that we have discussed here may well be a *proprium* of analytico-synthetic KOSs designed for specialized contexts or domains. After all, the epistemological profile of a general faceted classification might have less of a pragmatist inflection than Kaiser’s and Vickery’s systems for special libraries did, though only a detailed study of past attempts to design general faceted classifications can answer this. Nevertheless, our findings do suggest that any characterization of faceted KOSs as being rationalist *tout court* runs the risk of undervaluing the degree to which such systems actually exhibit epistemologico-methodological hybridity. Some commentators who consider faceted KOSs to embody primarily a rationalist approach have accepted that such systems have empirical and pragmatist dimensions to them, but have sought to minimize the latter by claiming that they are methodologically underdeveloped. With regard to Kaiser and Vickery, at least, one may wonder whether such a move is an analytically productive one. To be sure, their respective discussions of term collection and admonitions to attend to the interest of the user community for which a KOS was being constructed tended not to go into great detail: however, the very generality of their advice on these points allowed their statements to serve as broad and flexible guidelines for classifiers or indexers working in a variety of business-organizational or scientific-technological contexts and dealing with different subject domains without obligating them to follow specific methods for determining subject scope or selecting terms that might prove to be overly constraining in, or inappropriate for, certain contexts. At any rate, it seems more analytically fruitful to consider the interactions between rationalism, empiricism, and pragmatism in analytico-synthetic KOSs like Kaiser’s and Vickery’s than to shoehorn such systems within the category of rationalism alone.



## References

- Beghtol, C. (1986). Semantic validity: concepts of warrant in bibliographic classification systems. *Library Resources & Technical Services*, 30: 109–125.
- Broughton, V. (2011). In Memoriam: Brian Vickery, September 11, 1918–October 17, 2009. *Cataloging & Classification Quarterly*, 49: 42–46.
- Dousa, T. M. (2008). Empirical observation, rational structures, and pragmatist aims: epistemology and method in Julius Otto Kaiser's theory of systematic indexing. In C. Arsenault & J. T. Tennis (eds.), *Culture and identity in knowledge organization: Proceedings of the Tenth International ISKO Conference, 5–8 August 2008, Montréal, Canada* (Advances in Knowledge Organization, 11; pp. 242–248). Würzburg: Ergon Verlag.
- Fuller S., Hjørland B., Ibekwe-SanJuan, F., Ma, L., Mai, J.-E., Tennis, J., & Warner, J. (2013). Uncovering Epistemological Assumptions Underlying Research in Information Studies. Panel at the 76th Annual Meeting of the ASIS&T (Association for Information Science & Technology), Montréal, Canada, 1-5 November, 2013. Retrieved January 21, 2014, from: [http://www.asis.org/asist2013/restricted/ASIST\\_AM13/](http://www.asis.org/asist2013/restricted/ASIST_AM13/)
- Hjørland, B. (1998). The classification of psychology: a case study in the classification of a knowledge field. *Knowledge Organization*, 25: 162–201.
- Hjørland, B. (2002). Epistemology and the socio-cognitive perspective in information science. *Journal of the American Society for Information Science and Technology*, 53(4): 257–270.
- Hjørland, B. (2003). Fundamentals of knowledge organization. *Knowledge Organization*, 30: 87–111.
- Hjørland, B. (2011). The importance of theories of knowledge: indexing and information retrieval as an example. *Journal of the American Society for Information Science and Technology*, 62(1): 72–77.
- Hjørland, B. (2013a). Theories of knowledge organization—theories of knowledge. *Knowledge Organization*, 40: 169–181.
- Hjørland, B. (2013b). Facet analysis: the logical approach to knowledge organization. *Information Processing and Management*, 49: 545–557.
- Kaiser, J. (1911). *Systematic indexing*. London: Isaac Pitman & Sons.
- La Barre, K. (2010). Facet analysis. *Annual Review of Information Science and Technology*, 44: 243–284.
- Sales, R. de (2012). *A presença de Kaiser no quadro teórico do tratamento temático da informação (TTI)*. Doctoral Thesis, State University of São Paulo. Retrieved 17 January 2014, from: <http://www.bv.fapesp.br/pt/dissertacoes-teses/85422/a-presenca-de-kaiser-no-quadro-teorico-do-tratamento-tematic/>
- Svenonius, E. (2000). *The intellectual foundation of information organization*. Cambridge, MA/London: The MIT Press.
- Vickery, B. C. (1958). *Classification and indexing in science*. New York: Academic Press/London: Butterworths Scientific Publications.
- Vickery, B. C. (1960). *Faceted Classification: a guide to construction and use of special schemes*. London: Aslib.
- Vickery, B. C. (1966). *Faceted classification schemes* (Rutgers Series on Systems for the Intellectual Organization of Information, Vol. 5). New Brunswick, NJ: Graduate school of Library Service, Rutgers, the State University.